

FIG. 1

VAKTI-1 cDNA and its translation into  
 amino acid sequence

Frame 2

ATG	CAT	GGA	GTG	GAC	CTG	TAG	GCG	ACT	TGC	ATC	GTC	TTC	AAC	ATG	AAG	ATA	GCC
		10			19			28			37			46			55

M K I A  
I—MEMC-1—>  
I->HF6479

T	V	S	V	L	L	P	L	A	L	C	L	I	Q	D	A	A	S	I	K	N
ACA	GTG	TCA	GTG	CTT	CTG	CCC	TTG	GCT	CTT	TGC	CTC	ATA	CAA	GAT	GCT	GCC	AGT	AAG	AAT	
	64				73			82			91			100			109			

MEMC-1—>										CHEF-1—>									
E	D	Q	E	M	C	H	E	F	Q	A	F	M	K	N	G	K	L	F	C
GAA	GAT	CAG	GAA	ATG	TGC	CAT	GAA	TTT	CAG	GCA	TTT	ATG	AAA	AAT	GGA	AAA	CTG	TTC	TGT
	124				133			142			151			160			169		

<—CHEF-14—										<—CHEF-2—									
CHEF-11—>																			
P	Q	D	K	K	F	F	Q	S	L	D	G	I	M	F	I	N	K	C	A
CCC	CAG	GAT	AAG	AAA	TTT	TTT	CAA	AGT	CTT	GAT	GGA	ATA	ATG	TTC	ATC	AAT	AAA	TGT	GCC
	184				193			202			211			220			229		

HF6479<—I

T	C	K	M	I	L	E	K	E	A	K	S	Q	K	R	A	R	H	L	A
ACG	TGC	AAA	ATG	ATA	CTG	GAA	AAA	GAA	GCA	AAA	TCA	CAG	AAG	AGG	GCC	AGG	CAT	TTA	GCA
	244				253			262			271			280			289		

R	A	P	K	A	T	A	P	T	E	L	N	C	D	D	F	K	K	G	E
AGA	GCT	CCC	AAG	GCT	ACT	GCC	CCA	ACA	GAG	CTG	AAT	TGT	GAT	GAT	TTT	AAA	AAA	GGA	GAA
	304				313			322			331			340			349		

R	D	G	D	F	I	C	P	D	Y	Y	E	A	V	C	G	T	D	G	K
AGA	GAT	GGG	GAT	TTT	ATC	TGT	CCT	GAT	TAT	TAT	GAA	GCT	GTT	TCT	GGC	ACA	GAT	GGG	AAA
	364				373			382			391			400			409		

T	Y	D	N	R	C	A	L	C	A	E	N	A	K	T	G	S	Q	I	G
ACA	TAT	GAC	AAC	AGA	TGT	GCA	CTG	TGT	GCT	GAG	AAT	GCG	AAA	ACC	GGG	TCC	CAA	ATT	GGT
	424				433			442			451			460			469		

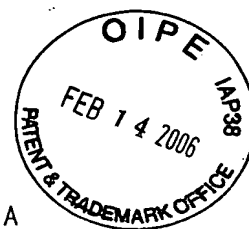
V	K	S	E	G	E	C	K	S	S	N	P	E	Q	V	R	S	I	V	S
GTA	AAA	AGT	GAA	GGG	GAA	TGT	AAG	AGC	AGT	AAT	CCA	GAG	CAG	GTG	AGG	TCA	ATT	GTC	AGC
	484				493			502			511			520			529		

L	M	G	N	T	G	R	L	T	S	N	S	K	STOP
CTG	ATG	GGA	AAT	ACT	GGG	AGG	CTA	ACT	TCA	AAT	AGT	AAG	TAG
	544				553			562			571		580

TTA	GGT	GGG	AGC	CTT	GGA	AGG	AAT	TAA	TTC	TTG	CTT	TAT	GTG	AAA	TGG	AAT	ACC	CAG	TTA
	604				613			622			631			640			649		

CTG	CCC	ACT	AAT	ATG	AAA	AAG	CTA	ATT	ATA	GTC	TCT	GAA	ACT	GGA	TCA	GAT	TAC	TTT	GGT
	664				673			682			691			700			709		

GGT	TAA	GAT	CTT	TCA	ATC	TAT	TGC	TGC	TTT	GTA	T
	724				733			742			749

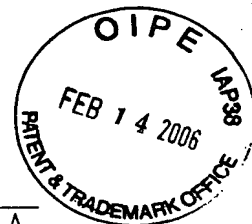


VAKTI-2 cDNA and its translation into  
amino acid sequence

FIG. 2A

Frame 2

ATG CAT GGA GTG GAG CTG TAG GCG ACT TGC ATC ATG TTG AAC ATG AAG ATA GCC  
10 19 28 37 46 55  
T V S V L L P L A L C L I Q D A A S I K N  
ACA GTG TCA GTG CTT CTG CCC TTG GCT CTT TGC CTC ATA CAA GAT GCT GCC AGT AAG AAT  
64 73 82 91 100 109  
Repeat 1  
E D Q E M \* C H E F Q M F M K N G K L F C  
GAA GAT CAG GAA ATG TGC CAT GAA TTT CAG GCA TTT ATG AAA AAT GGA AAA CTG TTC TGT  
124 133 142 151 160 169  
P Q D K K F F Q S L D G I M F I N K C A  
CCC CAG GAT AAG AAA TTT TTT CAA AGT CTT GAT GGA ATA ATG TTC ATC AAT AAA TGT GCC  
184 193 202 211 220 229  
HF 6479<—I  
T \* C K M I L E K E A K S Q K R A R H L A  
ACG TGC AAA ATG ATA CTG GAA AAA GAA GCA AAA TCA CAG AAG AGG GCC AGG CAT TTA GCA  
244 253 262 271 280 289  
typical Kazal domain  
R A P K A T A P T E L N \* C D D F K K G E  
AGA GCT CCC AAG GCT ACT GCC CCA ACA GAG CTG AAT TGT GAT GAT TTT AAA AAA GGA GAA  
304 313 322 331 340 349  
#  
R D G D F I C P D Y Y E A V C G T D G K  
AGA GAT GGG GAT TTT ATC TGT CCT GAT TAT TAT GAA GCT GTT TGT GGC ACA GAT GGG AAA  
364 373 382 391 400 409  
I  
T Y D N R C A L C A E N A K T G S Q I G  
ACA TAT GAC AAC AGA TGT GCA CTG TGT GCT GAG AAT GCG AAA ACC GGG TCC CAA ATT GGT  
424 433 442 451 460 469  
Repeat 2  
V K S E G E \* C K S S N P E Q D V \* C S A F  
GTA AAA AGT GAA GGG GAA TGT AAG AGC AGT AAT CCA GAG CAG GAT GTA TGC AGT GCT TTT  
484 493 502 511 520 529  
R P F V R N G R L G \* C T R E N D P V L G  
CGG CCC TTT GTT AGA AAT GGA AGA CTT GGA TGC ACA AGG GAA AAT GAT CCT GTT CTT GGT  
544 553 562 571 580 589  
#  
P D G K T H G N K C A M C A E L F L K E  
CCT GAT GGG AAG ACG CAT GGC AAT AAG TGT GCA ATG TGT GCT GAG CTG TTT TTA AAA GAA  
604 613 622 631 640 649  
A E N A K R E G E T R I R R N A E K D F  
GCT GAA AAT GCC AAG CGA GAG GGT GAA ACT AGA ATT CGA CGA AAT GCT GAA AAG GAT TTT  
664 673 682 691 700 709  
Repeat 3  
\* C K E Y E K Q V R N G R L F C T R E S D  
TGC AAG GAA TAT GAA AAA CAA GTG AGA AAT GGA AGG CTT TTT TGT ACA CGG GAG AGT GAT  
724 733 742 751 760 769  
#  
P V R G P D G R M H G N K C A L C A E I  
CCA GTC CGT GGC CCT GAC GGC AGG ATG CAT GGC AAC AAA TGT GCC CTG TGT GCT GAA ATT  
784 793 802 811 820 829  
F K R R F S E E N S K T D Q N L G K A E  
TTC AAG CGG CGT TTT TCA GAG GAA AAC AGT AAA ACA GAT CAA AAT TTG GGA AAA GCT GAA  
844 853 862 871 880 889



## FIG. 2B

Repeat 4

E	K	T	K	V	K	R	E	I	V	K	L	C	S	Q	Y	Q	N	Q	A	
GAA	AAA	ACT	AAA	GTT	AAA	AGA	GAA	ATT	GTC	AAA	CTC	TGC	AGT	CAA	TAT	CAA	AAT	CAG	GCA	
	904				913			922			931			940			949			
#																				
K	N	G	I	L	F	C	T	R	E	N	D	P	I	R	G	P	D	G	K	
AAG	AAT	GGA	ATA	CTT	TTC	TGT	ACC	AGA	GAA	AAT	GAC	CCT	ATT	CGT	GGT	CCA	GAT	GGG	AAA	
	964				973			982			991			1000			1009			
#																				
M	H	G	N	L	C	S	M	C	Q	V	Y	F	Q	A	E	N	E	E	M	
ATG	CAT	GGC	AAC	TTG	TGT	TCC	ATG	TGT	CAA	GTC	TAC	TTC	CAA	GCA	GAA	AAT	GAA	GAA	GCG	
	1024				1033			1042			1051			1060			1069			
#																				
I—>HF7665																				
K	K	A	E	A	R	A	R	N	K	R	I	E	S	G	K	A	T	S	Y	A
AAA	AAG	GCT	GAA	GCA	CGA	GCT	AGA	AAC	AAA	AGA	GAA	TCT	GGA	AAA	GCA	ACC	TCA	TAT	GCA	
	1084				1093			1102			1111			1120			1129			
#																				
Repeat 5																				
E	L	C	N	E	Y	R	K	L	V	R	N	G	K	L	A	C	T	R	E	
GAG	CTT	TGC	AAT	GAA	TAT	CGA	AAG	CTT	GTG	AGG	AAC	GGA	AAA	CTT	GCT	TGC	ACC	AGA	GAG	
	1144				1153			1162			1171			1180			1189			
#																				
N	D	P	I	Q	G	P	D	G	K	V	H	G	N	T	C	S	M	C	E	
AAC	GAT	CCT	ATT	CAG	GGC	CCA	GAT	GGG	AAA	GTG	CAC	GGC	AAC	ACC	TGC	TCC	ATG	TGT	GAG	
	1204				1213			1222			1231			1240			1249			
#																				
HF7665<—I																				
V	F	F	Q	A	E	E	E	E	K	K	K	K	E	G	E	S	R	N	I	K
GTT	TTT	TTC	CAA	GCA	GAA	GAA	GAA	GAA	AAG	AAA	AAG	AAG	GAA	GGC	GAA	TOA	AGA	AAC	AAA	
	1264				1273			1282			1291			1300			1309			
#																				
Repeat 6																				
R	Q	S	K	S	T	A	S	F	E	E	L	C	S	E	Y	R	K	S	R	
AGA	CAA	TCT	AAG	AGT	ACA	GCT	TCC	TTT	GAG	GAG	TTG	TGT	AGT	GAA	TAC	CGC	AAA	TCC	AGG	
	1324				1333			1342			1351			1360			1369			
#																				
K	N	G	R	L	F	C	T	R	E	N	D	P	I	Q	G	P	D	G	K	
AAA	AAC	GGA	CGG	CTT	TTT	TGC	ACC	AGA	GAG	AAT	GAC	CCC	ATC	CAG	GGC	CCA	GAT	GGG	AAA	
	1384				1393			1402			1411			1420			1429			
#																				
M	H	G	N	T	C	S	M	C	E	A	F	F	Q	Q	E	E	R	A	R	
ATG	CAT	GGC	AAC	ACC	TGC	TCC	ATG	TGT	GAG	GCC	TTC	TTT	CAA	CAA	GAA	GAA	AGA	GCA	AGA	
	1444				1453			1462			1471			1480			1489			
#																				
Repeat 7																				
A	K	A	K	R	E	A	A	K	E	I	C	S	E	F	R	D	Q	V	R	
GCA	AAG	GCT	AAA	AGA	GAA	GCT	GCA	AAG	GAA	ATC	TGC	AGT	GAA	TTT	CGG	GAC	CAA	GTG	AGG	
	1504				1513			1522			1531			1540			1549			
#																				
N	G	T	L	I	C	T	R	E	H	N	P	V	R	G	P	D	G	K	M	
AAT	GGA	ACA	CTT	ATA	TGC	ACC	AGG	GAG	CAT	AAT	CCT	GTG	CGT	GGA	CCA	GAT	GGC	AAA	ATG	
	1564				1573			1582			1591			1600			1609			
#																				
H	G	N	K	C	A	M	C	A	S	V	F	K	L	E	E	E	E	K	K	
CAT	GGA	AAC	AAG	TGT	GCC	ATG	TGT	GCC	AGT	GTG	TTC	AAA	CTT	GAA	GAA	GAA	GAG	AAG	AAA	
	1624				1633			1642			1651			1660			1669			
#																				
N	D	K	E	E	K	G	K	V	E	A	E	K	V	K	R	E	A	V	Q	
AAT	GAT	AAA	GAA	GAA	AAA	GGG	AAA	GTT	GAG	GCT	GAA	AAA	GTT	AAG	AGA	GAA	GCA	GTT	CAG	
	1684				1693			1702			1711			1720			1729			
#																				
Repeat 8																				
E	L	C	S	E	Y	R	H	Y	V	R	N	G	R	L	P	C	T	R	E	
GAG	CTG	TGC	AGT	GAA	TAT	CGT	CAT	TAT	GTC	AGG	AAT	GGA	CGA	CTC	CCC	TGT	ACC	AGA	GAC	
	1744				1753			1762			1771			1780			1789			



## FIG. 2C

N D P I E G L D G K I H G N T C S M C E  
AAT GAT CCT ATT GAG GGT CTA GAT GGG AAA ATC CAC GGC AAC ACC TGC TCC ATG TGT GAA  
1804 1813 1822 1831 1840 1849

A F F Q Q E A K E K E R A E P R A K V K  
GCC TTC TTC CAG CAA GAA GCA AAA GAA AAA GAA AGA GCT GAA CCC AGA GCA AAA GTC AAA  
1864 1873 1882 1891 1900 1909

Repeat 9

R E A E K E T C D E F R R L L Q N G K L  
AGA GAA GCT GAA AAG GAG ACA TGC GAT GAA TTT CGG AGA CTT TTG CAA AAT GGA AAA CTT  
1924 1933 1942 1951 1960 1969

F C T R E N D P V R G P D G K T H G N K  
TTC TGC ACA AGA GAA AAT GAT CCT GTG CGT GGC CCA GAT GGC AAG ACC CAT GGC AAC AAG  
1984 1993 2002 2011 2020 2029

C A M C K A V F Q K E N E E R K R K E E  
TGT GCC ATG TGT AAG GCA GTC TTC CAG AAA GAA AAT GAG GAA AGA AAG AGG AAA GAA GAG  
2044 2053 2062 2071 2080 2089

E D Q R N A A G H G S S G G G G G N T Q  
GAA GAT CAG AGA AAT GCT GCA GGA CAT GGT TCC AGT GGT GGT GGA GGA GGA AAC ACT CAG  
2104 2113 2122 2131 2140 2149

Repeat 10

D E C A E Y R E Q M K N G R L S C T R E  
GAC GAA TGT GCT GAG TAT CGG GAA CAA ATG AAA AAT GGA AGA CTC AGC TGT ACT CGG GAG  
2164 2173 2182 2191 2200 2209

S D P V R D A D G K S Y N N Q C T M C K  
AGT GAT CCT GTA CGT GAT GCT GAT GGC AAA TCG TAC AAC AAT CAG TGT ACC ATG TGT AAA  
2224 2233 2242 2251 2260 2269

A K L E R E A E R K N E Y S R S R S N G  
GCA AAA TTG GAA AGA GAA GCA GAG AGA AAA AAT GAG TAT TCT CGC TCC AGA TCA AAT GGG  
2284 2293 2302 2311 2320 2329

Repeat 11

T G S E S G K D T C D E F R S Q M K N G  
ACT GGA TCA GAA TCA GGG AAG GAT ACA TGT GAT GAG TTT AGA AGC CAA ATG AAA AAT GGA  
2344 2353 2362 2371 2380 2389

K L I C T R E S D P V R G P D G K T H G  
AAA CTT ATC TGC ACT CGA GAA AGT GAC CCT GTC CGG GGT CCA GAT GGC AAG ACA CAT GGT  
2404 2413 2422 2431 2440 2449

N K C T M C K E K L E M E A A E K K R K  
AAT AAG TGT ACT ATG TGT AAG GAA AAA CTG GAA AGG GAA GCA GCT GAA AAA AAA AGA AAG  
2464 2473 2482 2491 2500 2509

R M K T G A I Q E K G A I Q E K G A M T  
AGG ATG AAG ACA GGA GCA ATA CAG GAG AAA GGA GCA ATA CAG GAG AAA GGA GCA ATG ACA  
2524 2533 2542 2551 2560 2569

K R I C V V N F E A C R E M E S L S A P  
ATG AGG ATC TGT GTC GTC AAT TTC GAA GCA TGC AGA GAA ATG GAA AGC TTA TCT GCA CCA  
2584 2593 2602 2611 2620 2629



## FIG. 2D

E	K	I	T	L	F	E	A	H	M	A	R	C	T	S	I	N	V	L	C
GAG	AAA	ATA	ACC	CTG	TTC	GAG	GCC	CAT	ATG	GCA	AGA	TGC	ACA	TCA	ATA	AAT	GTG	CTA	TGT
	2644				2653			2662			2671			2680			2689		
V	R	A	S	L	I	E	K	L	M	K	E	K	R	K	M	K	R	N	Q
GTC	AGA	GCA	TCT	TTG	ATC	GAG	AAG	CTA	ATG	AAA	GAA	AAA	AGA	AAG	ATG	AAG	AGA	AAT	CAA
	2704				2713			2722			2731			2740			2749		
V	A	S	P	Q	I	M	Q	R	M	S	A	V	N	F	E	T	I	STOP	
GTA	GCA	AGC	CCT	CAA	ATA	ATG	CAA	AGG	ATG	AGT	GCA	GTG	AAT	TTC	GAA	ACT	ATA	TAA	GGA
	2764				2773			2782			2791			2800			2809		
ACA	ATG	AAC	TCA	TCT	GCC	CTA	GAG	AGA	ATG	ACC	CAG	TGC	ACG	GTG	CTG	ATG	GAA	AGT	TCT
	2824				2833			2842			2851			2860			2869		
ATA	CAA	ACA	AGT	GCT	CAC	TGT	GCA	GAG	CTG	TCT	TTC	TAA	CAG	AAG	CTT	TGG	AAA	GGG	CAA
	2884				2893			2902			2911			2920			2929		
AGC	TTC	AAG	AAA	AAC	CAT	CCC	ATG	TTA	GAG	CTT	CTC	AAG	AGG	AAG	ACA	GCC	CAG	ACT	CTT
	2944				2953			2962			2971			2980			2989		
TCA	GTT	CTC	TGG	ATT	CTG	AGA	TGT	GCA	AAG	ACT	ACC	GAG	TAT	TGC	CCA	GGA	TAG	GCT	ATC
	3004				3013			3022			3031			3040			3049		
TTT	GTC	CAA	AGG	ATT	TAA	ACC	CTG	TCT	GTG	GTG	ACG	ATG	GCC	AAA	CCT	ACA	ACA	ATC	CTT
	3064				3073			3082			3091			3100			3109		
GCA	TGC	TCT	GTC	ATG	AAA	ACC	TGA	TAC	GCC	AAA	CAA	ATA	CAC	ACA	TCC	GCA	GTA	CAG	GGA
	3124				3133			3142			3151			3160			3169		
AGT	GTG	AGG	AGA	GCA	GCA	CCC	CAG	GAA	CCA	CCG	CAG	CCA	GCA	TGC	CCC	CGT	TTG	ACG	AAT
	3184				3193			3202			3211			3220			3229		
GAC	AGG	AAG	ATT	GTT	GAA	AGC	CAT	GAG	GGA	AAA	AAT	AAA	CCC	CAG	TTT	TGA	ATC	ACC	TAC
	3244				3253			3262			3271			3280			3289		
CTT	CAC	CAT	CTG	TAT	ATA	CAA	AGA	ATT	TTT	CGG	AGC	TTG	TTT	TAT	TTG	CTA	TAG	AAA	ACA
	3304				3313			3322			3331			3340			3349		
ATA	CAG	AGC	TTT	TGG	GAA	TGG	AAT	CAC	TGA	TTT	TCA	GTC	TTT	TCC	ATT	TCT	TTC	CTC	CTA
	3364				3373			3382			3391			3400			3409		
GAA	TCT	GTG	ATC	TGA	GGG	TAT	AAA	GAC	ATT	TCC	ACC	AAG	TTT	GAG	CCC	TCA	AAA	TGT	CCT
	3424				3433			3442			3451			3460			3469		
polyadenylation signal																			
GAT	TAC	AAT	GCT	GTC	TGT	CCA	ACT	GCC	TGT	TCA	ATA	AAA	GTA	AAC	TCA	GCA	GAA	AA	...
	3484				3493			3502			3511			3520			3529		
..... poly(A) tail																			



**FIG. 3**

Trypsin inhibition

